

Formação de Regiões Homogêneas no Combate à Criminalidade Violenta de Minas Gerais utilizando técnica de agrupamento hierárquico entre 2004 e 2009.

**Paulo Rodrigues Homem<sup>1</sup>**

### *RESUMO*

“O Acordo de Resultados é um contrato de gestão no qual são definidos, pelo governo, os resultados esperados para cada área de atuação governamental, também chamada de Sistemas Operacionais. Um dos objetivos é: a) melhorar a qualidade e eficiência dos serviços públicos prestados à sociedade; b) facilitar o controle social sobre a atividade administrativa estadual” (Lei Estadual nº 17.600/08). A avaliação de resultados da política pública de combate à criminalidade violenta é premissa do Governador de Minas Gerais como contrapartida da arrecadação de impostos junto aos contribuintes tanto pessoas jurídicas quanto pessoas físicas. Em particular, a avaliação da política pública de integração das entidades Corpo de Bombeiros, Polícia Civil e Polícia Militar. A integração é percebida pela população através das Regiões Integradas de Segurança Pública (RISP) definidas pelo Sistema Integrado de Defesa Social Este trabalho usará a sub-rotina “Single” (Vizinho mais Próximo) do método de agrupamento hierárquico para encontrar grupos homogêneos.

Palavras-Chaves: Agrupamento Hierárquico; Acordo de Resultados.

### INTRODUÇÃO

A Região Integrada de Segurança Pública (RISP) é uma divisão geográfica, criada pelo Sistema Integrado de Defesa Social, para implantação e monitoramento da política de combate à criminalidade pertencente ao Plano Estadual de Segurança Pública preconizado pelo Governador de Minas Gerais. As entidades “Corpo de Bombeiros”, “Polícia Civil”, “Polícia Militar” designam representantes para toda a sede de região implantada, objetivando discutir e solucionar as demandas dos cidadãos contidos nos boletins de ocorrência registrados em sua área de atuação. Este trabalho tem como objetivo propor subgrupos homogêneos que possam ser usados para avaliação da política pública de região integrada. As áreas integradas de segurança pública estão apresentadas na Tabela 1 (página 2). A matriz de dados é composta de setenta e quatro observações mensais, entre os anos de 2004 e 2009, dos crimes violentos registrados pelas Regiões Integradas de Segurança Pública (RISP) que são as variáveis de estudo. O pacote estatístico “MINITAB”, versão 13, foi utilizado na análise.

### METODOLOGIA

A análise estatística de agrupamentos ou *Clusters Analysis* (CORRAR et al, 2007), citado por LUDEWIG et al (2009), é uma técnica e algoritmo multivariado cujo objetivo seria encontrar e separar objetos similares ou seja, tratando a idéia de semelhança entre grupos, procurando um esquema de classificação onde os grupos ou subgrupos com certo nível de semelhança. O princípio da análise de agrupamento consistiria em que cada observação de uma amostra multivariada corresponda a um ponto num espaço euclidiano multidimensional. Os processos de classificação resultariam em agrupar os pontos em conjuntos que evidenciaríamos aspectos marcantes da amostra. O resultado final seria um gráfico de esquemas hierárquicos denominados “dendograma” contendo uma síntese dos resultados, o que ocasionaria alguma perda de informação. Como é fácil o seu entendimento, a classificação do diagrama pode ser discutida amplamente sem conhecimentos específicos de estatística. Segundo (JOHNSON; WICHERN, 1982), citado por LUDEWIG, as questões vitais das técnicas de agrupamento seria a definição de similaridade ou dissimilaridade, que é denominada “coeficiente de semelhança”. Na medida de similaridade, quanto maiores os valores observados, mais parecidos seriam os objetos comparados. Na medida de dissimilaridade, quanto maiores os valores observados, menos parecidos seriam os

---

1 Paulohomem53@hotmail.com

objetos comparados. A maioria dos algoritmos de análise de agrupamento utiliza o conceito de distância (dissimilaridade), exigindo do usuário o esforço da transformação entre elas. A medida mais conhecida na indicação da proximidade entre dois objetos é a distância euclidiana  $d(A,B)$ :

- $d(A,B) = (\sum [x_i(A) - x_i(B)]^2)^{1/2}$  onde  $i = 1, 2, \dots, n$  (total de observações);
- $x_i(A)$  é o elemento da  $i$ -ésima coordenada do objeto A;
- $x_i(B)$  é o elemento da  $i$ -ésima coordenada do objeto B;

Na formação de agrupamento possuem duas idéias básicas: coesão interna dos dados e isolamento externo entre os grupos. Nas técnicas hierárquicas, os grupos seriam classificados em subgrupos produzindo uma árvore de classificação. (CORRAR et al, 2007). Segundo (TIMM, 2002), as medidas de proximidade mais conhecidas seriam: 1) o tamanho da distância (dissimilaridade) entre as observações; 2) o grau de associação (similaridade) entre elas. Dado dois objetos  $y_r$  e  $y_s$ , a medida de distância tem as seguintes propriedades matemáticas:

- 1)  $d^2 \geq 0$
- 2)  $d_{yr} = d_{ys}$  se somente se  $y_r = y_s$  ;
- 3)  $d_{yrs} = d_{ysr}$  tal que a distância é simétrica. (TIMM, 2002).

O pesquisador define “a priori” a similaridade ou a quantidade de grupos desejados. O diagrama “Dendograma” apresenta, a partir da decisão do pesquisador, o resultado das associações. A similaridade no diagrama de árvore é conceituada igual a “100” quando não existe agrupamento das variáveis estudadas. O método é válido com mínimo de dois grupos novos. A partir do primeiro grupo novo formado, a similaridade decresce e o processo é reaplicado até o agrupamento de todas as variáveis. A quantidade ideal de novos grupos (*clusters*) escolhida nesse estudo (Figura 1), foi efetuada pelo método “*Clusters Variable*”, utilizando a mínima distância entre os grupos para junção (“*Linkage*”) dos grupos iniciais. A matriz das distâncias é calculada pela expressão matemática:

- $d_{mj} = \min (d_{kj} - d_{kl})$  onde :
- $d_{kj}$  é a distância entre os agrupamentos  $k$  e  $j$ ;
- $d_{kl}$  é a distância entre os agrupamentos  $k$  e  $l$ ;

Tabela 1- Região/Cidade-Sede

RISP 1 - Belo Horizonte
RISP 2 - Contagem
RISP 3 - Vespasiano
RISP 4 - Juiz de Fora
RISP 5 - Uberaba
RISP 6 - Lavras
RISP 7 - Divinópolis
RISP 8 - Governador Valadares
RISP 9 - Uberlândia
RISP 10 - Patos de Minas
RISP 11 - Montes Claros
RISP 12 - Ipatinga
RISP 13 - Barbacena
RISP 14 - Curvelo
RISP 15 - Teófilo Otoni
RISP 16 - Unai
RISP 17 - Pouso Alegre
RISP 18 - Poços de Caldas

## RESULTADOS

Apresentamos a seguir o dendograma (Figura 1) das regiões integradas de segurança pública.

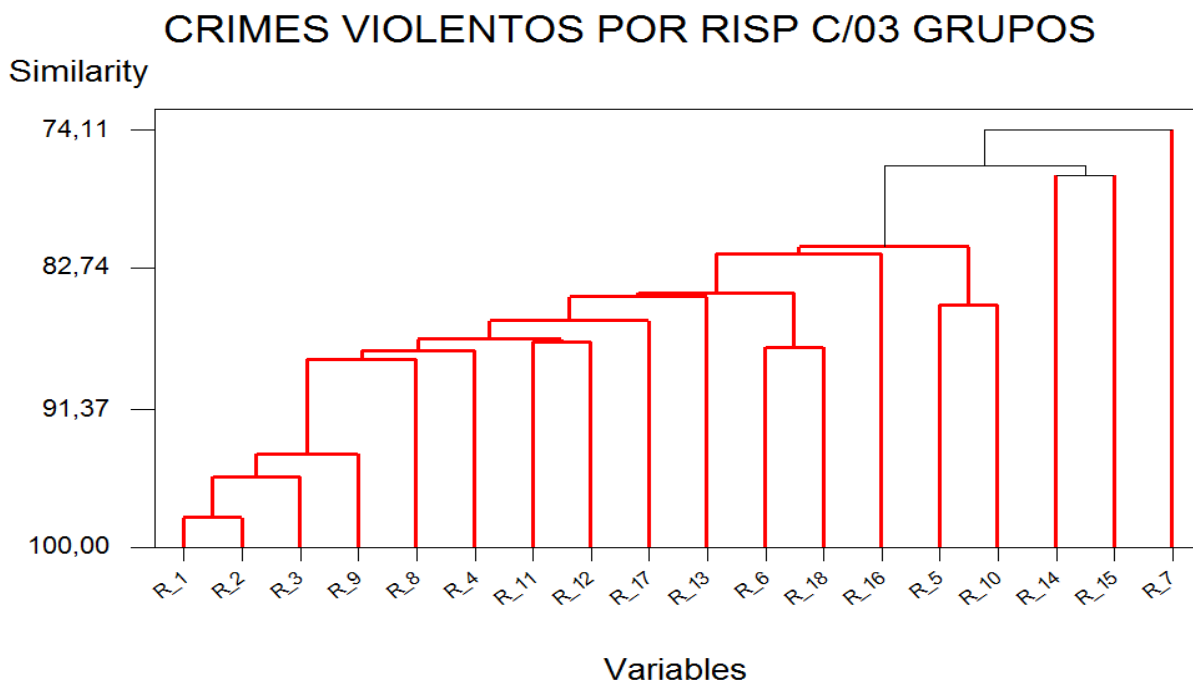


Tabela 2 – Nível de Similaridade e Grupos Formados pelo Método “Single”

Step	Nº Clusters	Similaridade Level	Cluster Joined/New	Step	Nº Clusters	Similaridade Level	Cluster Joined/New
0	18	100,00	Nenhum	9	9	85,95	1 e 17;1
1	17	98,17	1 e 2;1	10	8	84,96	5 e 10;5
2	16	95,65	1 e 3;1	11	7	84,47	1 e 13;1
3	15	94,22	1 e 9;1	12	6	84,24	1 e 6;1
4	14	88,35	1 e 8;1	13	5	81,79	1 e 16;1
5	13	87,80	1 e 4;1	14	4	81,37	1 e 5;1
6	12	87,62	6 e 18;6	15	3	76,94	14 e 15;14
7	11	87,26	11 e 12;11	16	2	76,32	1 e 14;1
8	10	87,11	1 e 11;1	17	1	74,11	1 e 7;1

Analisando o dendograma acima, podemos inferir que: 1) a região 7 (Divinópolis) possui a menor medida de similaridade (74,11) entre as regiões integradas; 2) as regiões 1 (BH) e 2 (Contagem) possuem a maior medida de similaridade (98,17) entre as regiões integradas; 3) as regiões 14 (Curvelo) e 15 (Teófilo Otoni); regiões 5 (Uberaba) e 10 (Patos de Minas); regiões 6 (Lavras) e 18 (Poços de Caldas) possuem a mesma medida de similaridade; 4) as regiões 11 (Montes Claros) e 12 (Ipatinga); regiões 8 (Gov. Valadares) e 4 (Juiz de Fora) possuem medida de similaridade muito próxima; 5) a região 16 (Unai) pode representar uma separatriz entre dois grandes grupos internos A (R\_5; R\_10; R\_14; R\_15; R\_7) e B (demais regiões). O mapa de Minas Gerais com os limites regionais está no anexo. Os grupos formados são: 1) grupo 3 (regiões de Curvelo e Teófilo Otoni); 2) grupo 2 (Divinópolis); 3) grupo 1 (Demais Regiões) com 76,94 de similaridade. As regiões com o mesmo nível de similaridade podem receber as mesmas metas no acordo de resultados.

## CONCLUSÃO

A análise por agrupamento hierárquico é muito eficiente na formação de grupos homogêneos com relação à criminalidade violenta entre 2004-09. Esta técnica pode ser usada no estabelecimento de metas proporcionais ao tamanho das regiões integradas de segurança pública e, por consequência, propor um acordo de resultados com distorções regionais mínimas.

## BIBLIOGRAFIA

1. BILODEAU, M. & BRENNER, D. (1999) – Theory of Multivariate Statistics, SPRINGER-VERLAG, N. Y. Inc.;
2. HÄRDLE, W. & HLAVKA, Z. (2007) – Multivariate Statistics: Exercises and Solutions, SPRINGER-VERLAG, N. Y. Inc.;
3. LUDEWIG, D. R et al, (2009) – O Processo de Gestão de Custos e Planejamento de Resultados utilizando técnicas de análise estatística de agrupamentos, Revista *Scientiarum Techonology*, Maringá (PR), volume 31 nº 2, p. 215-220;
4. MINITAB FOR WINDOWS (1998) Versão 13: Reference Manual, USA, *MINITAB INC.*;
5. TIMM, N. H. (2003) – Applied Multivariate Analysis, SPRINGER-VERLAG, N. Y. Inc.

ANEXO

